Video Compression Algorithm Based on Frame Difference Approaches

MuzhirShaban Al-Ani[1]andTalal Ali Hammouri[2]

Muzhir@gmail.comtalalah@yahoo.com

Abstract

The huge usage of digital multimedia via communications, wireless communications, Internet, Intranet and cellular mobile leads to incurable growth of data flow through these Media. The researchers go deep in developing efficient techniques in these fields such as compression of data, image and video. Recently, video compression techniques and their applications in many areas (educational, agriculture, medical …) cause this field to be one of the most interested fields.

Wavelet transform is an efficient method that can be used to perform an efficient compression technique. This work deals with the developing of an efficient video compression approach based on frames difference approaches that concentrated on the calculation of frame near distance (difference between frames). The selection of the meaningful frame depends on many factors such as compression performance, frame details, frame size and near distance between frames. Three different approaches are applied for removing the lowest frame difference. In this paper, many videos are tested to insure the efficiency of this technique, in addition a good performance results has been obtained.

Key Words:

Discrete Wavelet Transform, Video Compression, Frame Selection Key, Compression Performance and Near Distance Calculation.

1. Introduction

There are many methods used for video compression, and the most famous and apply technique is Moving Picture Experts Group (MPEG), which is an ISO/ITU standard for compressing digital video. MPEG performslossy compression for each frame similar to JPEG, which means pixels from the original images are permanently removed. The most famous methods are concentrated of Discrete Cosine Transform (DCT) and Discrete Wavelet Transform (DWT).

The grey scale image gives 256 levels of possible intensity for each pixel, so these images refer to 8 bits per pixel (bpp). The typical RGB color images, with 8 bits for Red, 8 bits for Green, and 8 bits for Blue, then the intensity I is defined by (I=R+G+B). The human eye is most sensitive to variations in intensity, so the most difficult part of compressing a color image lies in the compressing of the intensity.

Digital video consists of a stream of images captured at regular time intervals. The images are represented as digitized samples containing visual (color and intensity) information at each spatial and temporal location. Visual information at each sample point may be represented by the values of the three basic color components RGB color space. A video signal can be sampled in either frames (progressive) or fields (interlaced). In progressive video, a complete frame is sampled at each time instant. While an interlaced video only a half of the frame is captured (either odd or even rows of samples) at a particular time instant which are called fields. In this paper we deal with a sequence of complete frames in which we calculate the frame differencing between the consequences frames according to a specific threshold.

2. Literature Reviews

Many papers and researches are published related to this subject and some of these are listed below:

Chen Guanghua et al. proposed an AVS real time high definition video encoder for its high memory bandwidth and large calculation complexity caused by the new coding features of variable block size and 4-tap filter. The proposed design is implemented on FPGA with operating frequency of 150 MHz and can support 1080p (1920*1080)/30fps AVS real time encoder [1].

XueBai et al. introduced a new color model, Dynamic Color Flow, which incorporates motion estimation into color modeling in a probabilistic framework, and adaptively changes model parameters to match the local properties of the motion. The proposed model accurately and reliable describes changes in the science's appearance caused by motion across frames [2].

EsaRahtu et al. implemented a new salient object segmentation method based on combining a saliency measure with a conditional random field. The proposed saliency measure is formatted using a statistical framework and local feature contrast in illumination, color, and motion information [3].

Mohammed et al. presented an improved DC prediction mode based on the distance between the predicted and reference pixels. The proposed system adaptively selects the number of prediction mode for each 4*4 or 8*8 blocks, which reduces the number of overhead bits and computational cost of the encoder [4].

EugeniyBelyaev et al. proposed a new spatial scalable and low complexity video compression algorithm based on multiplication free three dimensional discrete pseudo cosine transform. This paper shows an efficient results compared with H.264/SVC as well as it can be used for robust video transmission over wireless channels [5].

Linh Tran Ho et al., introduced a robust motion estimation scheme using a direction integrated genetic algorithm to speed up the encoding process of H.264/AVC video compression as well as to keep low bits to code frames. The obtained results show that the proposed motion estimation obtains the expressive results for both the number of bits to code a frame and the time cost to code a frame [6].

EvgencyKaminsky et al., proposed an effective DCT-domain video encoder architecture that

decreases the computational complexity of conventional hybrid video encoders by reducing the number of transform operations between the pixel and DCT domain. The proposed system is based on the conventional hybrid coder and on a set of fast integer composition DCT transform. The proposed architecture may be used for the future Internet and 4G applications [7].

Cong Dao Han et al., implemented a novel search algorithm which utilizes an adaptive hexagon and small diamond search to enhance search speed. Simulation results showed that the proposed approach can speed up the search process with little effect on distortion performance compared with other adaptive approaches [8].

Jerome Gorin et al., produced an abstract decoder models (ADM) of MPEG decoders as programs described in a dataflow language, which naturally expresses potential parallelism between tasks of an application which makes an ADM description suitable for implementation to a wide variety of platforms for uni-processor systems to FPGAs [9].

Enrico Baccaglini et al., presented a performance compression between multiple description coding and unequal loss protection as tools to deliver multimedia data to mobile user. The compression is performed using as a case study JPEG 2000 coded images and H.264/AVC video sequences transmitted over lossy packet networks. The simulation results show that both schemes allocate the same amount of redundancy for any given encoding output rate to protect the transmitted information [10].

V. Vijayalakshmi et al., proposed a new video encryption scheme for sensitive applications. The objective of the proposed system is to analyze a secure and computational feasible video encryption algorithm for MPEG video to improve the security of existing algorithm by combining encryption in Intra and Inter frames and to test the algorithm against the common attacks [11].

Mohamed Haj Taieb et al., introduced a novel distributed video coding scheme with adaptive puncturing that sends more parity bits when it is not the case. This scheme is based on the argument that redirecting the party bits where they are the most effective, will improve the compression results. Simulation results demonstrate that lower bit rates are achieved [12].

Bibhas Chandra Dhara et al., presented a fast motion estimation algorithm and in residual frame coding proposed a fast method based on block truncation coding with pattern fitting concept is employed. The proposed video coding method is afast one with a good quality at the responsible bit rate [13].

ChangfengNiu et al., proposed a novel method for moving object segmentation in the H.264 compressed domain. The proposed method utilizes some characters of H.264 besides motion information with no more decoding required. Experimental results of several H.264 compressed video sequences demonstrate the good segmentation quality of the proposed approach [14].

Boxin Shi et al., introduced a new coding plus color correction scheme for multi-view video by exploring the color redundancy. Some advanced features in H.264 codec are utilized as automatic color annotation, and then an optimization based colorization is performed to render the color picture [15].

3. MPEG Types

MPEG is an asymmetrical system. It takes longer to compress the video than it does to decompress it in the DVD player, PC, set-top box or digital TV set.There are many types of MPEG such as [9, 11].

MPEG-1 (Video CDs)

Although MPEG-1 supports higher resolutions, it is typically coded at 352x240 x 30fps (NTSC) or 352x288 x 25fps (PAL/SECAM). Full 704x480 and 704x576 frames (BT.601) were scaled down for encoding and scaled up for playback [16]. MPEG-1 uses the YCbCr color space with 4:2:0 sampling, but did not provide a standard way of handling interlaced video. Data rates were limited to 1.8 Mbps, but often exceeded[16].

MPEG-2 (DVD, Digital TV)

MPEG-2 provides broadcast quality video with resolutions up to 1920x1080. It supports a variety of audio/video formats, including legacy TV, HDTV and five channel surround sound. MPEG-2 uses the YCbCr color space with 4:2:0, 4:2:2 and 4:4:4 sampling and supports interlaced video. Data rates are from 1.5 to 60 Mbps [16].

MPEG-4 (All Inclusive and Interactive)

MPEG-4 is an extremely comprehensive system for multimedia representation and distribution. Based on a variation of Apple's QuickTime file format, MPEG-4 offers a variety of compression options, including low-bandwidth formats for transmitting to wireless devices as well as high-bandwidth for studio processing.

A major feature of MPEG-4 is its ability to identify and deal with separate audio and video objects in the frame, which allows separate elements to be compressed more efficiently and dealt with independently. User-controlled interactive sequences that include audio, video, text, 2D and 3D objects and animations are all part of the MPEG-4 framework[16].

MPEG-7 (Meta-Data)

MPEG-7 is about describing multimedia objects and has nothing to do with compression. It provides a library of core description tools and an XML-based Description Definition Language (DDL) for extending the library with additional multimedia objects. Color, texture, shape and motion are examples of characteristics defined by MPEG-7[16].

MPEG-21 (Digital Rights Infrastructure)

MPEG-21 provides a comprehensive framework for storing, searching, accessing and protecting the copyrights of multimedia assets. It was designed to provide a standard for digital rights management as well as interoperability. MPEG-21 uses the "Digital Item" as a descriptor for all multimedia objects. Like MPEG-7, it does not deal with compression methods[16].

4. Two Dimensional DWT

Two dimensional DWT is obtained via the implementation of low pass and high pass filters on rows and columns of image respectively. A low pass filter and a high pass filter are chosen, such that they exactly halve the frequency range between themselves. This filter pair is called the Analysis Filter pair. First, the low pass filter is applied for each row of data, thereby getting the low frequency components of the row. But since the lpf is a half band filter, the output data contains frequencies only in the first half of the original frequency range. The high pass filter is applied for the same row of data, and similarly the high pass components are separated, and placed by the side of the low pass components [17]. This procedure is done for all rows. As mentioned above, the LL band at the highest level can be classified as most important, and the other 'detail' bands can be classified as of lesser importance, with the degree of importance decreasing from the top of the pyramid to the bands at the bottom [17]. DWT is a multispectral technique used for converting signal or image into four different bands such as low-low (LL), low-high (LH), high-low (HL) and high-high (HH) as demonstrated in figure (1)
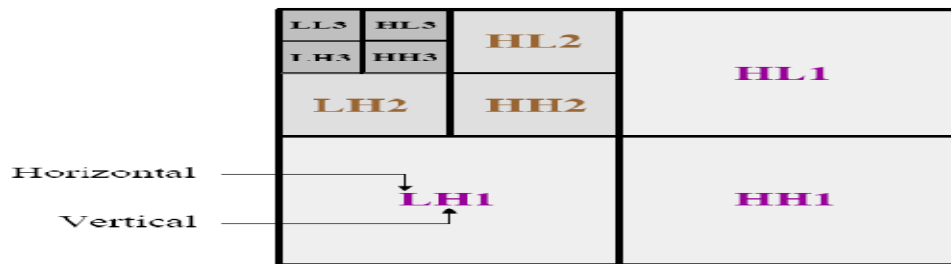

Figure (1) Decomposition of image applying DWT

5. Frame Difference Measure

Color histograms are frequently used to compare images because they are simple to compute, and tend to be robust regarding small changes in camera view point[18]. An image histogram H ( ) refers to the probability mass function of image intensities. Computationally, the color histogram is formed by counting the number of pixels belonging to each color. Usually a color quantization phase is performed on the original image in order to reduce the number of colors to consider in computing the histogram and thus the size of the histogram itself. There are a number of ways to compare color histograms. One simple method is the absolute difference between two color histograms as mentioned in the following equation.

$$\text{Diff } \mathbf{rgb}(I_i, I_j) = \sum_{K=1}^{n} (| H^r_i(k) - H^r_j(k) |$$
$$+ | H^g_i(k) - H^g_j(k) | + | H^b_i(k) - H^b_j(k)) \quad \text{...(1)}$$

Where, $H^r_i(j)$ is the histograms for color r in frame i.

6. System Implementation of Video Compression Approaches

The implemented system concentrated on frames difference approaches in which used for the compression technique. This system is implemented via many stages as illustrated in figure (2) and these stages are listed below:

- Filtering stage that generates the de-nosing video.
- Frame extraction that generates frames for a certain time.
- Frame selection that generates effective frames.
- Frame reordering that prepared frames.
- Applying 2D DWT stage that generates the LL-band.
- Video construction that construct the final compressed video.

The main core of this system is the frame difference approach, in which pass the similar frames and select the different frames depends on a certain specified threshold. In addition the selected frames are compressed via applying two dimensional discrete wavelet transform.
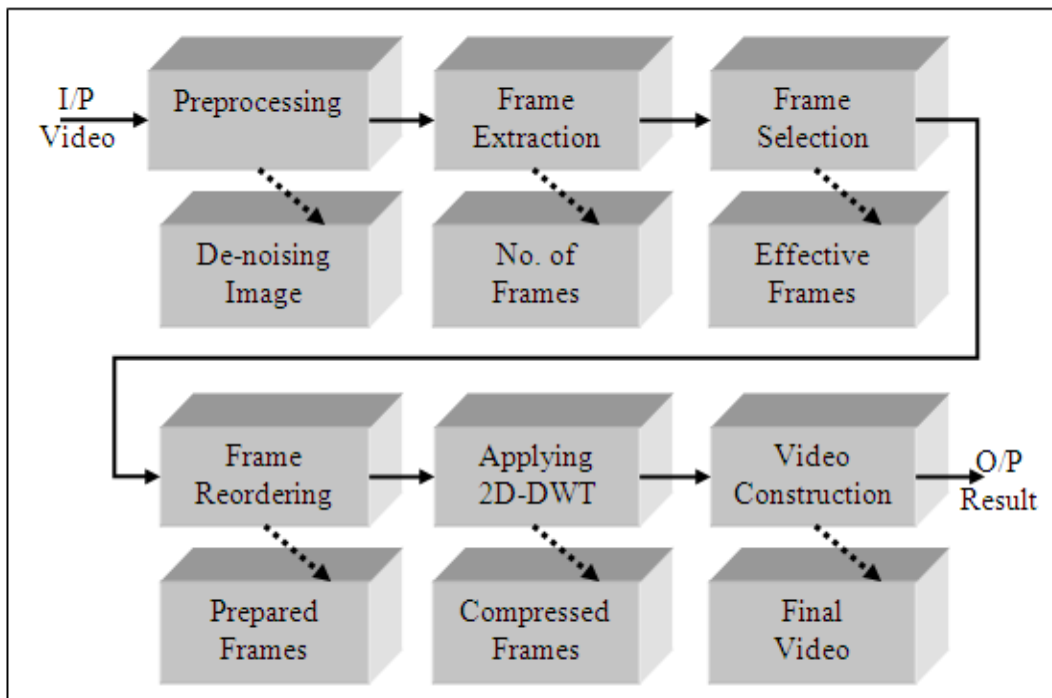

Figure (2) the proposed frame selectionkey approach

7. Frames Difference Approaches for Video Compression

Three different approaches are applied for removing the lowest frame difference; in these methods the similar frames are removed mainly. Three different methods are suggested to remove the similar frames in which frames difference approaches between each consecutive frame of the

extracted frames are applied. Summary of the structural diagram of the three approaches is shown in Figure (3).
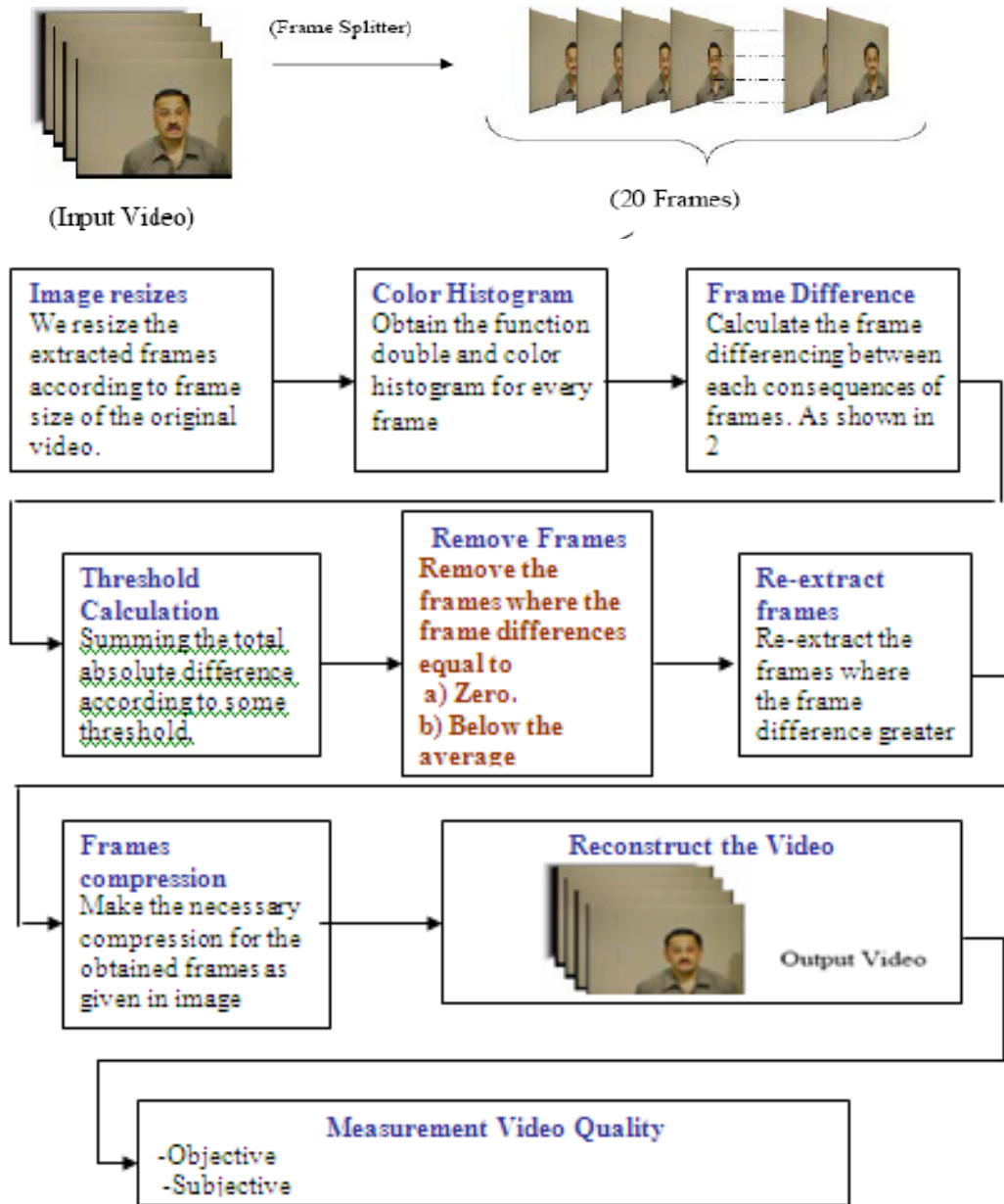

Figure (3) Structural Diagram of frames difference approach

a) **Zero DifferenceApproach**

In this approach frames are removed when the distance between any two consecutive frames is zero. It is clear in extracting video frames that many consecutive pair frames have zero difference, especially when the number of frame extracted above twenty frames per second. In

these cases it is better to remove that frames in which the frames difference is zero, this will minimize the number of frames to be re-extracted. Many of the frames difference between consecutive frames are zero; in this case such frames will be removed.

b) Mean Difference Approach

In this method, the mean value of the frames difference is calculated. Mathematically the mean (average) is obtained by dividing the sum of the observed values of frames difference by the number of observations. Then remove the frames where the frames difference between any consecutive frames is lower than the mean value of the frames difference.

c) Percentage Difference Approach

In this approach different types of videos are examined according to frame details, frame size and the obtained frames differences. This approach concerns mainly in removing the lowest frames differenceaccording to some   specific percentage. The percentage depends upon many factors such as compression performance, frame details, frame size and near distance between frames

8. Results and Analysis

Many types of videos are tested; these videos are extracted to forms different sets of frames according to number of frames per second. We have tested different types of videosare tested according to:

-Video types.
- Wavelet families' i.e. haar, db2, sym2, coif2, bior2.2, and rbio2.2.
- Levels of decomposition i.e. 1, 2, 3 and 4.
-Frames extraction i.e. frames per second 10, 15,   20 & 25.

Different sizes of frames where examined. Then frames differences are measured between each two nearest frames and this procedure continue up to the end of frames. Results of the mentioned approaches were illustrated in the following details

a) Zero Difference Approach

 Figures 6a & 7a represent the frames difference including zero's difference for frames per second 15 and 20 respectively. Part b of the mentioned figures 6b & 7b represents frames difference excluding zero's difference frames, while the removed frames in which the frames difference equal to zero are shown in figures 6c & 7c. As increasing extracting frames per second 10, 15, 20 and 25, the frames difference equal to zero not necessarily increased as shown in the figures 6c and 7c, since the frames differences in general decreased as the frames per second increased but not exactly equal to zero.
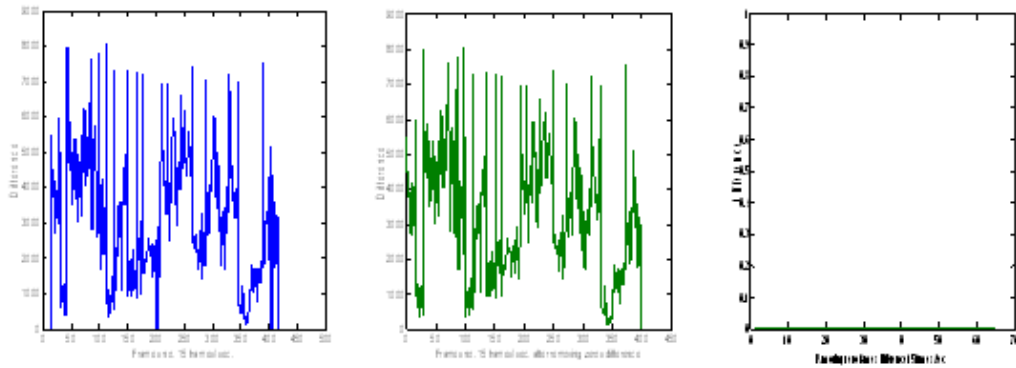
Figure (6) a) Zero Difference Approach.(b) Frames difference after removing
(15frames/sec)   zero difference   c) theremoved zero
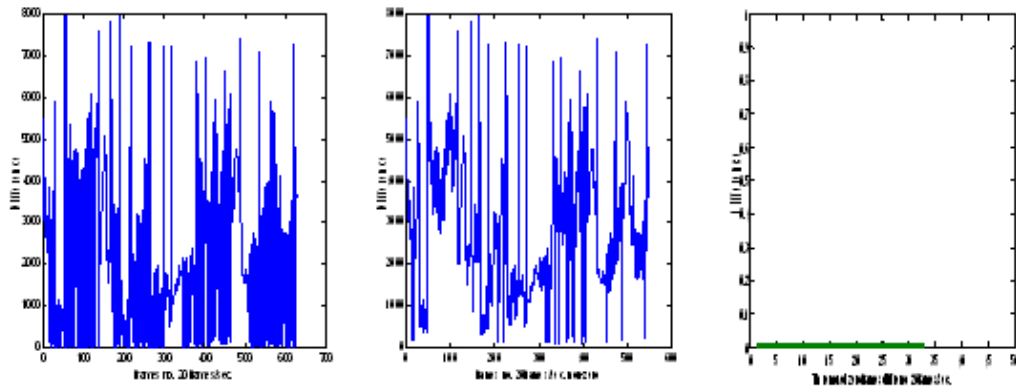difference

Figure (7) a) Zero Difference Approach,(b) Frames difference after removing   c) theremoved zero difference
(20frames/sec)   zero difference

b) Mean difference approach

Figures 8a & 9a represent the frames difference including mean's difference for frames per second 15 and 20 respectively. Part b of the figures 8b & 9b represents frames difference including frames below the mean of the overall difference frames,the green line represents the mean of the frames differences. The frames difference above the green line will bemaintained and compressed, while the frames difference below the green line will be removed. Part c of the figures 8c, 8c & 9c represents the frames to be removed. As increasing extracting frames per second, the frames difference decreased and the frames removed are larger
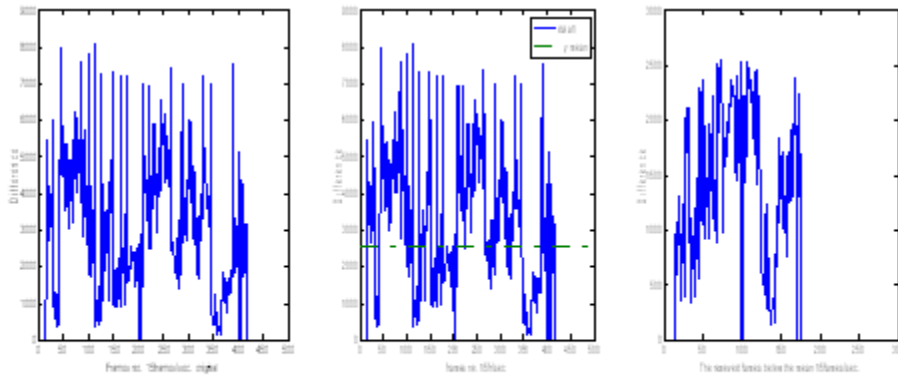
Figure (8) a) Zero Difference Approach.(b) Frames difference after removing zero difference   c) theremoved zero difference
(15frames/sec)
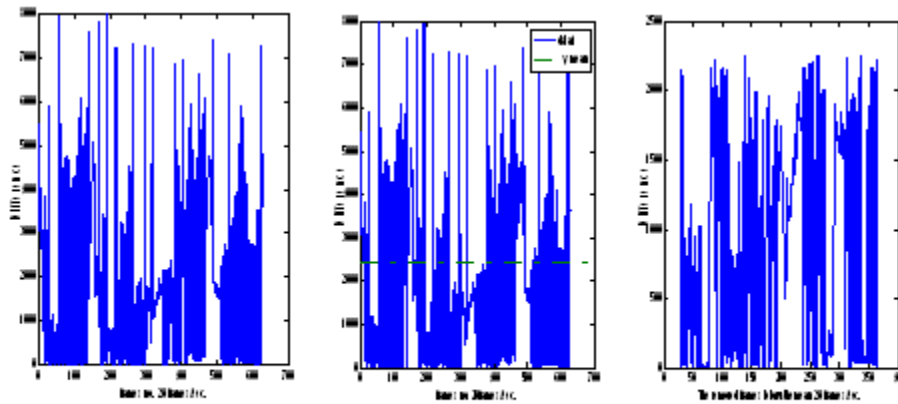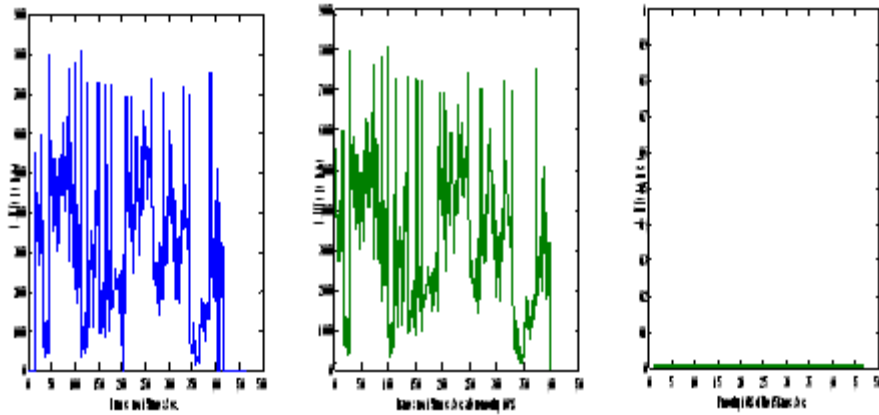

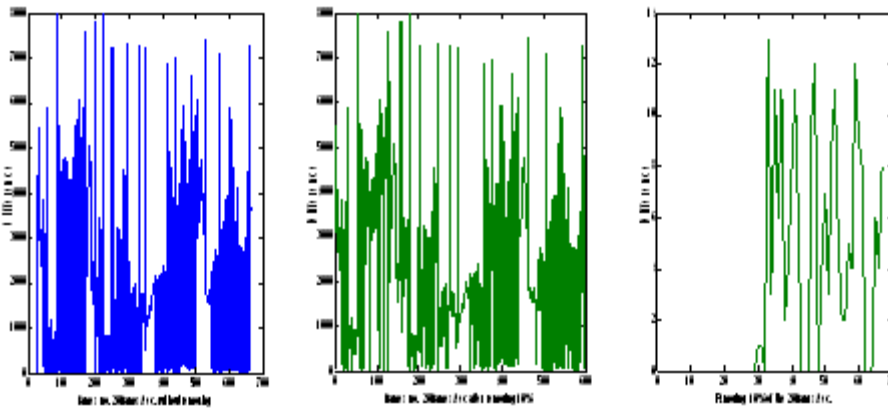Figure (9) a) Zero Difference Approach.(b) Frames difference after removing zero difference   c) theremoved zero difference
(15frames/sec)

C) Percentage Difference Approach

The implemented approach concerns mainly in removing the lowest frames difference.Figures 10a & 11a represent the frames difference for different frames per second 15, & 20 respectively. Part b offigures 10b & 11b represents frames difference after removing 10% of the lowest frames difference. Part c of the figures 10c & 11c represents the frames to be removed. Also different percentages like 15%, 20% were removed of the frames difference.

Figure (10) a) Mean Difference Approach.(b) Frames difference after removing
(15frames/sec)   10%lowest difference   c) theremoved 10%
difference

Figure (11) a) Mean Difference Approach.(b) Frames difference after removing
(20frames/sec)   10%lowest difference   c) theremoved 10%
difference

9. Subjective Video Quality   Assessment

To evaluate the dimension of the overall video quality,   a special   simple questionnaire was used. Examples of questions were asked after the presentation of the original and the compressed video, results of video quality assessment are given intable (1) below. From the obtained results it is clear that video quality increased with the increasing of frames per second. Evaluations of 25 frames give maximum video quality scores comparing with the   remaining numbers of frames per second (10, 15, and 20).

Table (1) Video quality a ssessment (zero difference approaches).

Video Quality Factors	10 frames	15 frames	20 frames	25 frames
How would you rate video colors?	3.9	4.0	4.2	4.4
How would you rate video contrast?	4.0	4.2	4.3	4.5
How would you rate video borders?	4.0	4.1	4.4	4.6
How would you rate the movement continuity?	3.8	4.0	4.2	4.4
Did you notice any flicker in the sequence?	annoying	not annoying	not annoying	not annoying
Did you notice any smearing in the sequence?	annoying	little annoying	not annoying	not annoying

10. Conclusion

An efficient video compression approach based on frames difference approaches are developed that concentrated on the calculation of frame near distance. Many factors are applied in the selection of meaningful frames, in which eliminate the similar frames. The implemented system passes into many steps; preprocessing, frame extraction, frameselection, frame reordering, 2D-DWT, then video construction. Different types of videos are introduced to test the system. The output compressed video is in a good quality and good performance as well as it has a specific compression ratio.

References

[1]   Chen Guanghua et al., "VLSI Implementation of sub-pixel interpolator for AVS encoder", LSMS/ICSEE 2010, part II, LNCS 6329, pp 351-359, 2010. (fulltext2)

[2]   XueBai et al. "Dynamic Color Flow: A Motion Adaptive Color Model for Object Segmentation in Video", ECCV 2010, part V, LNCS 6315, pp 617-630, 2010. (fulltext4)

[3]   EsaRahtu et al., "Segmenting Salient Objects from Images and Videos", ECCV 2010, part V, LNCS 6315, pp 366-379, 2010. (fulltext5)

[4]   Mohammed et al., "Enhanced Intra Coding of H.264/AVC Advanced Video Coding Standard with Adaptive Number of Modes", ATM 2010, LNCS 6335, pp 361-372, 2010. (fulltext11)

[5]   EugeniyBelyaev et al., "Scalable Video Coding Based on Three Dimensional Discrete Pseudo Cosine Transform", ruSMART/NEW2AN 2010, LNCS 6294, pp 448-459, 2010. (fulltext12)

[6]   Linh Tran Ho et al., "Direction Integrated Genetic Algorithm for Motion Estimation in H.264/AVC", ICIC 2010, LNAI 6216, pp 279-286, 2010. (fulltext18)

[7]   EvgencyKaminsky et al., "DCT-domain coder for Digital Video Applications", Journal of Real Time Image Processing, 14 August 2010. (fulltext21)

[8]   Cong Dao Han et al., "An Adaptive Fast Search Algorithm for Block Motion Estimation in H.264", Journal of Zhejiang University Science C (Computer & Electron), Vol. 11, No. 8, pp 637-644, 2010. (fulltext30)

[9]   Jerome Gorin et al., "LLVM – Based and Scalable MPEG-RVC Decoder", Journal of Real Time Image Processing, 27 July 2010. (fulltext32)

[10]   Enrico Baccaglini et al., "Image and Video Transmission: A Comparison Study of Using Unequal Loss Protection and Multiple Description Coding", Multimedia Tools Applications, 27 July 2010. (fulltext33)

[11]   V. Vijayalakshmi et al., "Efficient of Intra and Inter Frames in MPEG Video", CNSA 2010, CCIS 89, pp 93-104, 2010. (fulltext34).

[12]   Mohamed Haj Taieb et al., "Turbo Code Using Adaptive Puncturing for Pixel Domain Distributed Video Coding", ICISP 2010, LNCS 6134, pp 342-350, 2010. (fulltext42)

[13]   Bibhas Chandra Dhara et al., "A video Coding Technique Using Octagonal Motion Search and BTC-PF Method for Fast Reconstruction", AST/UCMA/ISA/CAN 2010, LNCS 6059, 480-490, 2010. (fulltext44)

[14]   ChangfengNiu et al., "Moving Object Segmentation in the H.264 Compressed Domain", ACCV 2009, part II, LNCS 5995, pp 45-54, 2010. (fulltext59)

[15]   Boxin Shi et al., " Color Correction and Compression for Multi-View Video Using H.264 Features", ACCV 2009, part III, LNCS 5996, pp 43-52, 2010. (fulltext60)

[16]   MPEG Definition from PC Magazine Encyclopedia
       http://www.pcmag.com/encyclopedia_term/0%2C2542%2Ct%3DMPEG&i%3D47295%2C00.asp

[17]   Satish Kumar. S, Copyright (c) 2010, website: http://www.debugmode.com/imagecmp/classify.htm

[18]   GianluigiCiocca et al "Dynamic story boards for video content summarization" UniversitàdegliStudi di Milano-Bicocca.

Author

Muzhir Shaban Al-Anihas received Ph. D. in Computer & Communication Engineering Technology, ETSII, Valladolid University, Spain, 1994. Assistant of Dean at Al-Anbar Technical Institute (1985). Head of Electrical Department at Al-Anbar Technical Institute, Iraq (1985-1988), Head of Computer and Software Engineering Department at Al-Mustansyria University, Iraq (1997-2001), Dean of Computer Science (CS) & Information System (IS) faculty at University of Technology, Iraq (2001-2003). He joined in 15 September 2003 Electrical and Computer Engineering Department, College of Engineering, AppliedScienceUniversity, Amman, Jordan, as Associated Professor. He joined in 15 September 2005 Management Information System Department, AmmanArabUniversity, Amman, Jordan, as Associated Professor, then he joined computer science department in 15 September 2008 at the same university.